ABSTRACT OF THE DISCLOSURE

The present invention is directed to a method for speech-based information retrieval in Mandarin Chinese, considering a monosyllabic structure of the Chinese language, and a whole class of syllable-based indexing terms, including overlapping segments of syllables and syllable pairs separated by a few syllables. The strong discriminating capabilities of such syllable-based indexing terms have been verified. Special approaches for better utilizing such capabilities, including fusion with the word- and character-level information and improved approaches to obtain better syllable-based features and query expressions and so on, are disclosed too.